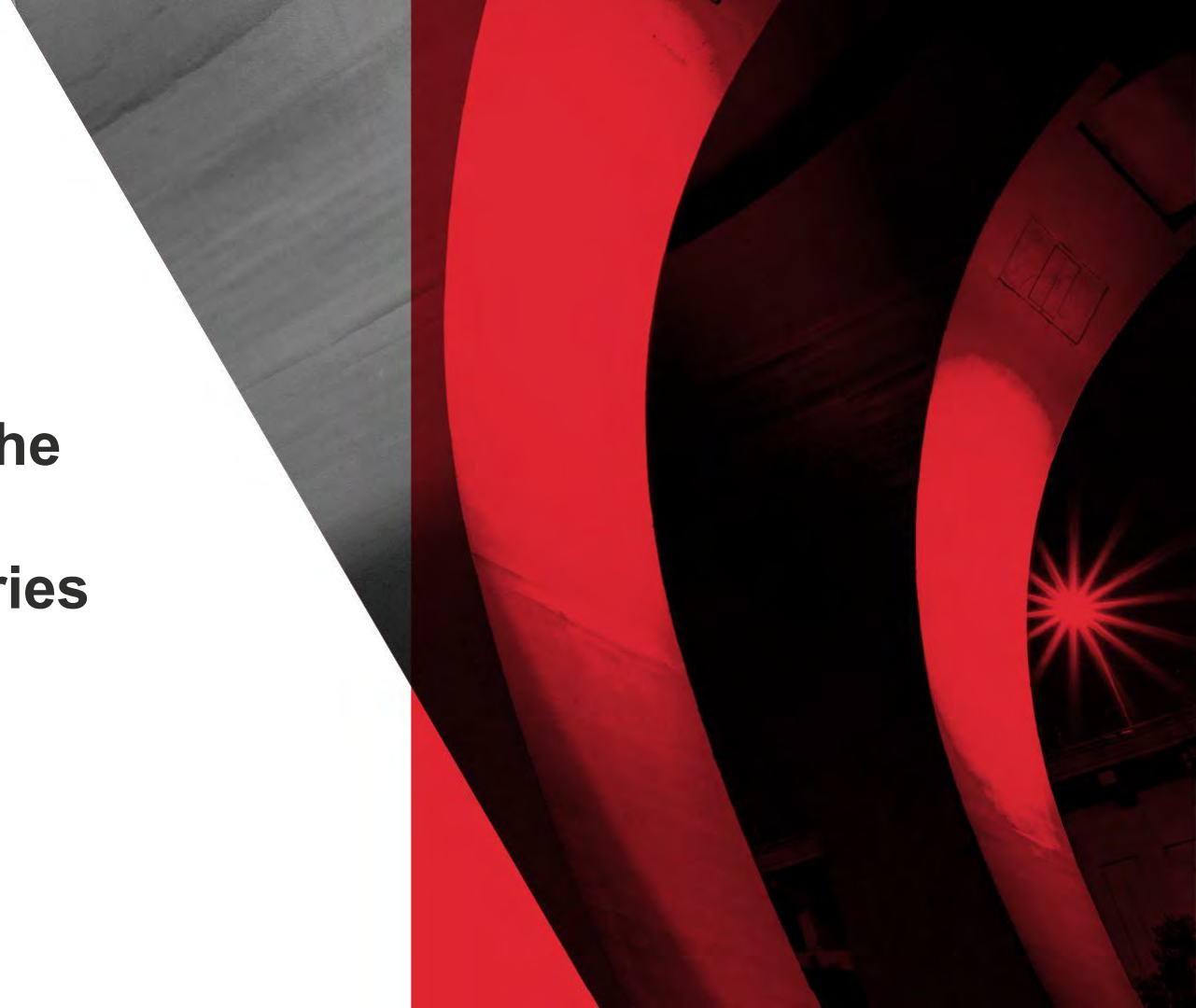
2023 INTERNET2 TECHNOLOGY exchange

Entity Categories: Federation to Scale

Pål Axelsson, Sunet / Swedish Research Council
Joanne Boomer, University of Missouri
Judith Bush, OCLC
Scott Cantor, Ohio State University, The
Steven Premeau, University of Maine System
Albert Wu, Internet2

Pål Axelsson, SUNET/Swedish Research Council
Joanne Boomer, University of Missouri
Judith Bush, OCLC
Scott Cantor, The University of Ohio
Steven Premeau, University of Maine
Albert Wu, Internet2

Introducing the * Access Entity Categories



The identity federation dilemma

Services need information about the user to be able to let them log in, but the Identity Providers want to be in control of which personal data Service Providers get access to

- Common is that Identity Providers doesn't release any attributes to unknown Service Providers
- This prevents users from logging into the services they need to access to complete their research or studies
- Users who are prevented from logging in often do not contact the helpdesk/support, instead they try to find other, less secure and non-organized ways into the service
- The tool for minimizing the dilemma is Entity Categories

What is Entity Categories?

Entity Categories = “group federation entities that share common criteria. The intent is that all entities in a given entity category are obliged to conform to the characteristics set out in the definition of that category.”

REFEDS Entity Categories

- Research & Scholarship (R&S)
- Anonymous Access (v2 released)
- Pseudonymous Access (v2 released)
- Personalized Access (v2 released)
- Data Protection Code of Conduct (CoCo v2, regional for EU/EEA)
- Hide from Discovery

Federation Operator Responsibility

Federation Registrar/Operator checks when assigning a REFEDS Entity Category to a Service Provider:

- The Service Provider has requested assignment of the Entity Category and complies with this entity category's registration criteria.
- Each Entity Category has a specific set of registration criteria
- The Service Provider's request to be assigned the Entity Category has been reviewed against the provided REFEDS Guidelines and approved by the federation registrar.

Entity Support Categories

Identity Providers may indicate support for specific Entity Categories to facilitate discovery and improve the user experience at Service Providers. Self-assertion is the typical approach used, but this is not the only acceptable method.

Support Entity Categories is also a good way to measure Identity Provider adoption of Entity Categories.

Note that Support Entity Category the technical marking in metadata is not the same as for Entity Categories.

REFEDS Anonymous Access Category

“Candidates for the Anonymous Access Entity Category are Service Providers that offer a level of service based on proof of successful authentication.”

This Entity Category does not support personalization for a specific user between different logged in session. Two non-personal attributes is included.

Specification and identifier: https://refeds.org/category/anonymous

REFEDS Pseudonymous Access Category

“Candidates for the Pseudonymous Access Entity Category are Service Providers that offer a level of service based on proof of successful authentication and offer personalization based on a pseudonymous user identifier. The Service Provider must be able to effectively demonstrate this need to their federation registrar (normally the Service Provider’s home federation) and demonstrate their compliance with regulatory requirements concerning personal data through a published Privacy Notice.”

Specification and identifier: https://refeds.org/category/pseudonymous

REFEDS Personalized Access Category

“Candidates for the Personalized Entity Category are Service Providers that have a proven need to receive a small set of personally identifiable information about their users in order to effectively provide their service to the user or to enable the user to signal their identity to other users within the service. The Service Provider must be able to effectively demonstrate this need to their federation registrar (normally the Service Provider’s home federation) and demonstrate their compliance with regulatory requirements concerning personal data through a published Privacy Notice.”

The successor of REFEDS Research and Scholarship Entity Category.

Specification and identifier: https://refeds.org/category/personalized

Using the * Access Entity Categories in the InCommon Federation

Community Feedback Draft

https://docs.google.com/document/d/1B45F1GKHjUY0j3QNIQ_XojFziKN9W02xCuL49vdAQRk/

Entity categories and their attributes

User Attribute	Personalized	Pseudonymous	Anonymous
user identifier (subject-id)	✅	🚫	🚫
pseudonymous pairwise user identifier (pairwise-id)	🚫	✅	🚫
person name (displayName, givenName, sn)	✅	🚫	🚫
email address (mail)	✅	🚫	🚫
organization (schacHomeOrganization)	✅	✅	✅
affiliation (eduPersonScopedAffiliation)	✅	✅	✅
assurance (eduPersonAssurance)	✅	✅	🚫

Entity categories and implementation decisions

Implementer	Personalized	Pseudonymous	Anonymous
InCommon	R&S terms for reg	Open registration	Open registration
IdP	Consider supporting all as entity categories; use all as templates for attribute release in bilateral agreements with SPs.		
	schacHomeOrganization & scopes in eduPersonAffiliation		
	New subject identifier	New pairwise identifier	
		Consider restricted eduPersonScopedAffiliation release.	
SP	Pick one for registration. Use these as templates for attribute release in bilateral agreements with IdPs		
	Understand different signals of affiliation		
	New subject identifier	New pairwise identifier	
	Note institutional interpretation of eduPersonScopedAffiliation “roles” may differ.		

InCommon's Attribute Release Recommendations

User Attribute	Personalized	Pseudonymous	Anonymous
user identifier (subject-id)	✅	🚫	🚫
pseudonymous pairwise user identifier (pairwise-id)	🚫	✅	🚫
person name (displayName, givenName, sn)	✅	🚫	🚫
email address (mail)	✅	🚫	🚫
organization (schacHomeOrganization)	✅	✅	✅
affiliation (eduPersonScopedAffiliation)	✅	✅ *	✅ *
assurance (eduPersonAssurance)	✅	✅	🚫

Legend

✅ Required by category

* See discussions in Working with Required Attributes

🚫 Not allowed in category

Affiliation (eduPersonScopedAffiliation)

eduPersonScopedAffiliation - https://wiki.refeds.org/display/STAN/eduPerson

Specifies a person's affiliation(s) within a particular domain within an organization. Left component uses controlled vocabulary @ right component is the administrative domain to which the affiliation applies. Multi-valued attribute

For Privacy Preserving Purposes InCommon Recommends releasing

ONLY **member** and **affiliate** affiliations

for both **Anonymous** and **Pseudonymous** Entity Categories

“Member” is intended to include faculty, staff, student and other persons with a full set of basic privileges that go with membership in the university community.

“Affiliate” indicates the holder has some definable affiliation to the university NOT captured by faculty, staff, student, employee, alum and/or member. (ex. Volunteers, parents of students, guests)

Affiliation (eduPersonScopedAffiliation)

Affiliation != Authorization

Do not assume these affiliations directly translate to authorization access to any service.

Each organization determines the precise interpretation of the affiliation values.

Service Providers, if your access policy is compatible then eduPersonScopedAffiliation is simple and scalable way to enable access (e.g. any member of an organization, as defined by that organization, can access my service).

If you need more information to determine access or authorization, then the entity categories do not fit your situation – eduPersonEntitlement is likely a good use for individualized service needs.

Organization (schacHomeOrganization)

schacHomeOrganization - https://wiki.refeds.org/display/STAN/SCHAC+Releases

Specifies a person's home organization using the domain name of the organization.

- Issuers of schacHomeOrganization attribute values via SAML are strongly encouraged to publish matching shibmd:Scope elements as part of their IDP's SAML metadata.
- Relaying Parties receiving schacHomeOrganization values via SAML are strongly encouraged to check attribute values against the Issuer's published shibmd:Scope elements in SAML metadata, and may discard any non-matching values.

Single Value Attribute

Because schacHomeOrganization is single valued, this could pose an issue for an IdP that represents multiple organizations (e.g. a system-wide IdP representing multiple universities in a system) or SPs that assume schacHomeOrganization will fit their need as it cannot convey complex organization relationships.

Name (sn, givenName, displayName)

- By this point of the presentation… there are no easy answers
  (and many "correct" options).
- The (traditionally) Western concepts of "family" and "given" names can represented in the "sn" and "givenName" attributes.
  - This representation quickly becomes challenging – if not impossible – when trying to represent names from other cultures.
- The displayName attribute provides a way for the full name to be represented in the desired format.
  - Cultural differences make it "impossible" to utilize this attribute to determine name parts.
- As institutions embrace diversity, equity, and inclusion, the provided name is less likely to be a legal name.
  - University of Maine System just revised their policies to release preferred name information unless legal name is explicitly required.

Characteristics of the new user identifiers

Common characteristics:

- Long-lived and non-reassignable; scoped; specified case, characters, length:

```
<value> = <uniqueID> "@" <scope>
<uniqueID> = (ALPHA / DIGIT) 0*126(ALPHA / DIGIT / "=" / "-")
<scope> = (ALPHA / DIGIT) 0*126(ALPHA / DIGIT / "-" / ".")
```

samlSubjectID (subject-id)

- Omni-directional: globally-unique external key
- Value for a given subject is independent of the relying party to whom it is given

samlPairwiseID (pairwise-id)

- Uni-directional: provider specific unique external key
- Value depends upon the Service Provider : inhibits correlation attacks

Comparison to Older Identifiers

subject-id vs. eduPersonPrincipalName (or mail)

- Stability - EPPN tends to be name-based and subject to churn
- Added guarantee of non-reassignment
- Added length constraint

subject-id vs. eduPersonUniqueId

- Case insensitivity to match “typical” application behavior
- Added character set and length constraints

pairwise-id vs. SAML 2.0 persistent NameID (eduPersonTargetedID)

- Case insensitivity to match “typical” application behavior
- Added character set and length constraints
- Formatted as simple value@domain with domain (scope) checking
- Insulated from entityID/issuer changes

Next … and your questions

Be alert for a review of InCommon TAC's guidance document(s).

Those attending ACAMP, we welcome session proposals to discuss these issues and others…

Identifier migration?

Affiliation?

Authorization?

Migration Strategies

User Attribute	IdP	SP
user identifier (subject-id)	migrate	migrate
pseudonymous pairwise user identifier (pairwise-id)	migrate	migrate
person name (displayName, givenName, sn)	implement	
email address (mail)	Adapt for anon and pseudon	
organization (schacHomeOrganization)	Implement if not already implemented	May identify specific affiliation when IdP supports multiple institutions
affiliation (eduPersonScopedAffiliation)	Consider offering differentleft parts depending on the profile	May identify multiple specific affiliations when IdP supports multiple institutions and the subject is affiliated with multiple ins
assurance (eduPersonAssurance)	support	Consider whether more appropriate

References

https://refeds.org/category/personalized
https://refeds.org/category/anonymous
https://refeds.org/category/pseudonymous
https://docs.oasis-open.org/security/saml-subject-id-attr/v1.0/cs01/saml-subject-id-attr-v1.0-cs01.html
https://wiki.refeds.org/display/STAN/SCHAC+Releases
https://wiki.refeds.org/display/STAN/eduPerson